

Lifestyle surveys – the complete answer?

Regina Dengler, Heather Roberts, Lesley Rushton

Abstract

Study objectives – These were as follows: to study incompleteness of data, herein called item non-response, generated by a self completion questionnaire; to identify the characteristics of item non-responders and the types of questions liable to high item non-response rates; and to discuss possible reasons for item non-response.

Design – Item non-response patterns in 12 307 responders (62%) to a representative postal survey based on a stratified sample drawn from family health services authorities' (FHSA) registers were investigated.

Main outcome measures – Data were analysed for item non-response in three groups depending on when the questionnaire was returned (wave analysis). The overall completion rate of the questionnaire was examined and the natural logarithm of the proportion of completed questions was used as an outcome variable in multiple regression analysis. Item non-response to key questions and questions of different types was examined.

Results – Wave analysis: the overall completion rate of the questionnaire was 86% in questionnaires returned before the first reminder and 83%–84% in those sent back after subsequent reminders. Overall pattern of item non-response: respondents failed to complete a mean of 15% and a median of 10% of the questionnaire. All questions in the questionnaire had some item non-response, ranging from 1% to 85%. Completion rates were associated with gender, age, indicators of lower socioeconomic status, and general health status. Individual questions: particular types of questions were liable to have higher item non-response, for example, linked binary questions.

Conclusions – Item non-response in population postal surveys is likely to present problems in the interpretation of data by introducing bias additional to that of total non-response. Item non-response does not increase greatly with later returns, suggesting that the quality of data across responses generated by two reminders is similar. There are obstacles to reducing item non-response, such as respondent error or socioeconomic and health characteristics of the general population, that cannot be totally overcome. However, the evidence that individuals tend to complete only options within questions that apply to them and their positive behaviour is useful information for those designing

questionnaires and interpreting survey data.

(*J Epidemiol Community Health* 1997;51:46-51)

In postal studies, bias may be introduced through respondents failing to complete all the questions, referred to here as "item non-response".¹ This bias is additional to any bias caused by total non-response,² such as the under representation of young men^{3,4} and smokers.^{5,6} Therefore, a better understanding of item non-response should help in the interpretation of data generated by postal surveys. Reasons for item non-response vary and include poor question design, questionnaire layout, and the inability or unwillingness of the respondent to supply information.⁷

A previous study has focused on item non-response in a series of three postal surveys of patients who had recently undergone three different surgical procedures (each involving around 300 patients).⁸ In all three surveys, more than half of all respondents missed one or more items (56%, 63%, and 60%). The amount of variance in item non-response attributable to respondents' characteristics, such as gender, age, education, or health perception, in regression modelling was fairly low ($R^2 = 7\%$, 12% , and 9%), suggesting that patient characteristics do not relate strongly to item non-response. In all three surveys questions at the end of the questionnaire were no more likely to be omitted than those at the beginning and the two most frequently omitted questions were on family income (13%, 14%, and 23%) and employment status (6%, 12%, and 17%). The question on sexual satisfaction, a sensitive issue, was omitted by fewer respondents than expected by the authors (3%, 5%, and 11%).

This paper focuses on item non-response in 12 307 responders to a postal survey. It examines the characteristics of item non-responders, the types of questions particularly liable to a high item non-response rate, and explores possible reasons for item non-response.

Method

THE STUDY SAMPLE

Altogether 21 603 adults aged 16–70 years were selected in relation to health district in Trent Regional Health Authority from family health services authorities' (FHSA) registers for inclusion in the Trent health lifestyle survey. This postal survey generated 12 307 (62%) responses after two reminders and allowing for 1800 (8%) exclusions, mainly due to mail that could not be delivered.⁴ All respondents to the

Department of Public Health Medicine and Epidemiology, University Hospital, Queen's Medical Centre, Nottingham NG7 2UH
R Dengler
H Roberts
L Rushton

Correspondence to:
Dr H Roberts.

Accepted for publication
May 1996

survey were included as the study sample. Data were pre-coded, entered on computer, and verified. Extensive testing and screening of raw data were carried out, including consistency checking, frequencies, and cross tabulations.

THE QUESTIONNAIRE

The questionnaire was 14 pages long and contained mostly closed questions relevant to Trent's Regional Strategic Framework for Health Promotion,⁹ covering topics such as smoking, drinking, stress, health checks, and sociodemographic details. Respondents were asked to tick boxes or write numbers in the spaces provided. In this paper one question is referred to as an item.

PROCEDURE FOR ANALYSING ITEM NON-RESPONSE

The study was carried out in three parts. The first part was an investigation of item non-response for early and late responders (wave analysis). Each question on the questionnaire was analysed for three groups as follows:

- 1 Those who responded before the first reminder,
- 2 Those who responded between the first and the second reminder, and
- 3 Those who responded between the second reminder and the end of the data collection period.

Multiple regression analysis was then used to examine the overall completion rate of the questionnaire. The overall completion rate was measured by the proportion of questions answered out of the number of applicable questions for each individual. Since this distribution was skewed, the natural logarithm of the proportion was used as an outcome variable. Common sociodemographic characteristics of the sample, such as gender (male/female), council tenancy (yes/no), age (16–54/55–70), usage of car (no/yes), and school leaving age (aged 14 or younger/15 or older) as well as long term illness (no/yes) and self reported health status (good/fairly good or poor) were used as explanatory variables.

The third part concentrates on a number of selected questions chosen from the questionnaire. For some key questions^{9,10} comparison was made between item non-responders and all respondents in terms of the variables described in the previous paragraph. For other questions the pattern of item non-response was examined. Questions were chosen by the following criteria:

- 1 Key questions: smoking and alcohol consumption were chosen because of their relevance to *Health of the Nation* targets.¹⁰
- 2 Style of question.
 - a Smoking was also an example of a question which required the respondent to choose 1 out of 4 mutually exclusive options.
 - b Health checks and consumption of butter/margarine were used as examples of ordered category responses. These two sets of questions were chosen particularly

because they contained at least one question with high item non-response.

- c The set of questions on stress was chosen as a random example of binary (yes/no) responses. Other binary questions were also examined, but the results, which appear to follow a similar pattern, are quoted in less detail.

The questions examined are displayed in figure 1.

Results

CHARACTERISTICS OF RESPONDENTS

Forty three per cent of the respondents were male and 57% female. Generally, males aged 16–34 years were under represented in response, a common characteristic of this type of survey.¹¹ Of all respondents, 45% were in manual and 39% were in non-manual occupations.¹² In addition, 6% of respondents were currently unemployed. Compared with census information, respondents without a car and those living in council accommodation were under represented, suggesting that a greater proportion of non-responders were in lower income groups.⁴

WAVE ANALYSIS

Sixty eight per cent of all returned questionnaires (8315/12 307) were sent back before the first reminder, 17% (2101/12 307) between the first and second reminder, and the outstanding 15% (1891/12 307) after the final reminder. The mean completion rate of the questionnaire was 86% in questionnaires returned before the first reminder, 84% in those sent back after the first, and 83% sent back after the second reminder. There was little difference in item non-response for individual questions except in response to the question about smoking status, where item non-response actually decreased with each wave from 10% to 7%.

OVERALL PATTERN OF ITEM NON-RESPONSE

Overall, respondents failed to complete a mean of 15% of the questionnaire (median 10%, range 0% to 96%). All the questions in the questionnaire had some item non-response, ranging from less than 1% for a question about alcohol consumption to 85% in respondents who failed to indicate "other" non-smoking areas they would like to see, apart from four areas given previously in that set of questions. Item non-response was also no higher at the end of the questionnaire than at the beginning of it.

Respondents who omitted one or more of the explanatory variables had lower overall completion rates of the whole questionnaire. For example, respondents who omitted their gender completed, on average, 65% of the questionnaire. Generally, women had higher levels of item non-response than men (table 1). Item non-response also increased as school leaving age decreased, with tenure of council housing, and with having no use of a car. In

2 FOOD

This part is about how often you eat the foods listed below.
Please tick one box for every item

I eat –

	less than once a week or never	once a week or more – not most days	every or most days
butter or “hard” margarine or ghee	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>
low fat spreads	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>
polyunsaturated spreads	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>

(Please check the packet if you are not sure.)

4 ALCOHOL

Whether or not you drink alcohol, please answer the questions.

I drink alcohol –
Tick one box only.

1 ☐ regularly – once a week or more

2 ☐ occasionally – not every week

3 ☐

5 TOBACCO

Whether you are a smoker, an ex-smoker or have never smoked, please answer these questions.

Still thinking about smoking –
Tick one box only. You may need to write numbers in the spaces.

1 ☐ I've never smoked or have only smoked once or twice – go on to Part 6: Stress.

2 ☐ I stopped smoking years months ago – go on to Part 6: Stress.

3 ☐ I now smoke daily

4 ☐ I now smoke occasionally

6 STRESS

Stress can make people ill.
This part is about the things that make you anxious or worried.

I feel stressed by –
Tick one box for each item.

	no	yes	no	yes
my work or unemployment	1 <input type="checkbox"/>	2 <input type="checkbox"/>	relationships with other people	1 <input type="checkbox"/>
how much money I have	1 <input type="checkbox"/>	2 <input type="checkbox"/>	caring for my family	1 <input type="checkbox"/>
fear of unemployment	1 <input type="checkbox"/>	2 <input type="checkbox"/>	pain or illnesses I have	1 <input type="checkbox"/>
racial harassment	1 <input type="checkbox"/>	2 <input type="checkbox"/>	fear of illness	1 <input type="checkbox"/>
sexual harassment	1 <input type="checkbox"/>	2 <input type="checkbox"/>	other things –	1 <input type="checkbox"/>
housing	1 <input type="checkbox"/>	2 <input type="checkbox"/>	please say what,	2 <input type="checkbox"/>
my local environment (e.g. litter, dog dirt)	1 <input type="checkbox"/>	2 <input type="checkbox"/>	(e.g. changes in my life)	2 <input type="checkbox"/>
fear of crime	1 <input type="checkbox"/>	2 <input type="checkbox"/>
danger on the roads	1 <input type="checkbox"/>	2 <input type="checkbox"/>		

8 ABOUT YOUR HEALTH

The next questions are about five health checks which you may have had.

Some family doctors offer patients a special “well woman” or “well man” check up.
This usually includes height and weight checks, blood pressure and urine tests, heart and lung checks and advice about diet, exercise, alcohol and smoking.

Not all people want – or need – all these checks.

The last time I had these health checks was –
Tick one box for each item.

	under a year ago	1–3 years ago	4–5 years ago	over 5 years ago	never
eye test	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
dental check	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
hearing test	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
“well woman” or “well man” check	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
women only – cervical smear	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>

Figure 1 Questions analysed from the 1992 Trent health lifestyle survey.

Table 1 Estimated proportion of item non-response in relation to characteristics of respondents

	Proportion item non-response	Lower 95% confidence interval	Upper 95% confidence interval
Reference*	0.91		
Female	0.89	0.88	0.89
Aged 55–70 y	0.87	0.86	0.88
Left school before the age of 15 y	0.84	0.82	0.87
Lives in council accommodation	0.88	0.87	0.90
No usage of car	0.89	0.87	0.90
Self-reported health status fair or poor	0.90	0.89	0.90

*Reference: male, aged 16–54 years, left school aged 15 or older, lives in owner occupied, privately rented or “other” accommodation, has the use of a car, self reported health status “good”.

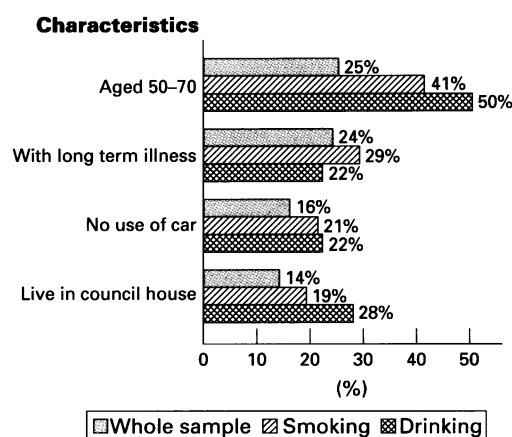


Figure 2 Characteristics of respondents who omitted the question on smoking status or alcohol consumption compared with the characteristics of the whole population. Of 12 307 respondents, 1153 left out the question on smoking status and 103 omitted the question on alcohol consumption.

addition, item non-responders were more likely to be in the 55–70 age range and classify their health as “not good”.

ITEM NON-RESPONSE TO SPECIFIC QUESTIONS

Smoking

Item non-response to the question about smoking status was 9%. Compared with all respondents, those who failed to complete the

smoking question were more likely to be aged 55–70 years, live in council-rented accommodation, and have a long term health problem/disability (fig 2). They were less likely to have the use of a car.

Alcohol consumption

The initial question on drinking status – that is, whether the respondent drank alcohol regularly, occasionally, or never, was omitted by only 1% of respondents. Twenty seven per cent of those who failed to complete the question were retired compared with 10% of the study sample. Consequently, a higher proportion of item non-responders to the question was aged 55–70 years. Non-responders were also more likely to be council tenants.

Health checks

Two per cent of respondents failed to answer all the questions on health checks. Figure 3 shows that questions about specific health checks were not answered by 3% to 12% of respondents, with the exception of cervical smears where 35% of women did not answer the question, although their demographic characteristics did not differ from other women in the study sample. Analysis of item non-responders to the question about their last visit to the dentist showed them to be: more females than males (65% versus 57%); older (56% versus 25% respondents aged 55–70); report their health as fairly or not good (52% versus 41%); to be less likely to have the use of a car (25% versus 16%) and more likely to be council tenants (23% versus 14%). A similar pattern was observed among item non-responders to questions about eye tests, hearing tests and well woman checks.

Item non-responders to questions about health checks were more likely to omit other parts of the questionnaire. For example, those who omitted the dental check question also missed questions on long term illness (37% compared with 6%), smoking (21% compared with 9%) and/or alcohol they had drunk in the week before the survey (29% compared with 17%).

Diet: fats and spreads

Three separate questions were asked about the frequency of consumption of “butter and hard margarine”, “low fat spreads” and “polyunsaturated spreads” (fig 1). Item non-response was about 25% on individual items.

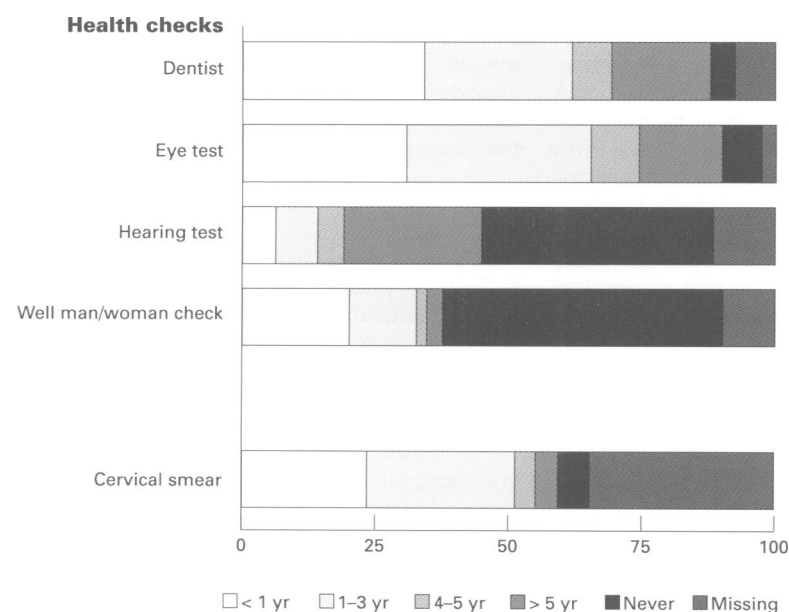


Figure 3 Health checks in 12 307 respondents (6830 women).

Only 1% of respondents omitted all three questions, and, of the 41% omitting either one or two of the questions, 84% reported that they ate one of the three fats/spreads "on every or most days".

Questions incorporating a series of "no"/"yes" (binary) questions

Binary questions required respondents to tick "yes" if an item from a predetermined set of options applied to them or "no" if it was not applicable. Some of the individual questions within these sets of questions had a very high item non-response rate. For example, the set of questions about stress offered 14 "yes"/"no" options. Non-response to the whole set of questions was 2% but item non-response within the set ranged from 12% to 50%. There was a consistently high incidence of item non-response (44% to 85%) when a set of questions ended with a request for open ended information. Such items asked, first, for a tick to indicate "yes" or "no" to the option of "other", followed by "please say what" and a line on which to write.

Discussion

The 61% response rate compares well with other postal lifestyle surveys using the same methodology. Response has been shown to vary between 50% and 60% but has been as high as 80%.¹³

Overall, item non-response was somewhat higher than in Guadagnoli and Cleary's study, though the questionnaire was about the same length. This was perhaps due to the different nature of their study and their analysis of only selected (core) questions. However, it is interesting that item non-response for individual questions was in the same order of magnitude in both studies, and for both studies the positioning of a question at the beginning or the end of the questionnaire made no difference.

The amount of variation in the proportion of completed questions explained by the respondents' gender and age, the only characteristics known before to the survey, was within 8%, about the same as in Guadagnoli and Cleary's study ($R^2=7\%$, 12% , and 9%), although they were also able to use other socio-demographic and health related characteristics.

If the underlying reasons for total non-response and item non-response were the same, then non-responders and item non-responders would be expected to share characteristics. For example, total non-response under represents males, younger people, those who are less educated and in lower socioeconomic groups, smokers, more frequent consumers of alcohol, and those who are ill. In this survey item non-responders did mirror some characteristics of total non-responders in that they tended to be less educated as assessed by school leaving age, poorer as indicated by tenancy and use of car, and less well according to two measures of self reported health status. Consequently, the potential for bias, already introduced by total

non-responders, is increased. However, the higher proportion of item non-responders who were more likely to be women or in the 55–70 age group may counteract the bias introduced by total non-response.

Data do not support the hypothesis that late responders omit more questions than early respondents as, although this was the case for the overall completion rate of the questionnaire, difference in item non-response for individual questions was small. Late responders were actually slightly more likely to complete the potentially sensitive question about smoking status.

The reasons for some item non-response may have been embedded in the questionnaire design, such as confusing layout of questions, poor instructions, and text difficulty. Some is likely to be due to mistakes, such as turning two pages over together or failing to go back to a question despite intending to do so. Item non-response may have been caused by characteristics of responders who have, for example, poor eyesight, lower levels of education or who are not familiar with terms such as "cervical" or "well woman check". The study shows that item non-response seems to cluster around specific types of questions, for example, sets of questions within which there are optional elements, such as the stress question noted, and where there are a series of linked questions, as in the example of fats and spreads. The pattern of responses suggests that respondents tend to tick options that apply to them, leaving all other options blank. For example, it may be that if stressed by money respondents tick "yes", but if unstressed by "other" factors they will fail to tick "no". Similarly, if someone eats butter often but rarely eats polyunsaturated margarine they will tend to indicate what they do eat and fail to indicate what they do not eat. If this is the case, then analysing positive responses only is justifiable. However, a better solution would be to redesign the set of questions so that each option has only one box which needs to be ticked to indicate "yes", offering an additional "none of these" option to check whether the set of questions has been missed.¹⁴ This has indeed been done in later lifestyle surveys, and generally item non-response has been reduced; for example the question on stress had 2% item non-response.

CONCLUSIONS

Item non-response in population postal surveys is likely to present problems in interpretation of data, introducing bias which, in part, exacerbates that introduced by total non-response. Unfortunately, the issue is not clear cut in that some aspects of item non-response also create biases that appear to counteract those caused by total non-response. It is encouraging that item non-response does not increase greatly with later returns, suggesting that the completeness of data across responses generated by reminders is similar.

There are obstacles to reducing item non-response that cannot be overcome, such as respondent error or socioeconomic and health

characteristics of the general population. However, the evidence that where questions have several options, individuals tend to complete only those that apply to them and their positive behaviour, is useful information for those designing questionnaires and interpreting survey data.

The authors gratefully acknowledge the support of Trent Regional Health Authority and the contribution of colleagues in the Department of Public Health Medicine and Epidemiology, in particular that of Katherine Fielding.

- 1 Moser CA, Kalton G. *Survey methods in social investigation*. 2nd ed. Aldershot: Dartmouth Publishing Company Limited, 1971.
- 2 Cartwright A. *Health surveys in practice and potential*. London: King's Fund Publishing Office, 1983.
- 3 Scott C. Research on mail surveys. *J R Stat Soc* 1961; 124(A):143–205.
- 4 Dengler R, Rushton L, Roberts HR, Magowan R. Results from a lifestyle survey: Trent Health. *HER* 1994;9:285–96.
- 5 Smith C, Nutbeam D. Assessing non-response bias: a case study from the 1985 Welsh Heart Survey. *HER* 1990;5: 381–6.
- 6 Roberts H, Bali B, Rushton L. Non-responders to a lifestyle survey: a study using telephone interviews. *Journal of the Institute of Health Education* 1996; 34: 57–61.
- 7 Gillies P. Questionnaires and interviews – notes on methodological considerations. In: Fraser RC, Goeting LM, editors. *Current approaches. Research methods in general practice*. Southampton: Duphar Laboratories Limited, 1991:39–50.
- 8 Guadagnoli E, Cleary PD. Age-related item non-response in surveys of recently discharged patients. *J Gerontol* 1992; 47:206–12.
- 9 Trent Regional Health Authority. *Strategic framework for health promotion in Trent region*. Trent Regional Health Authority; Sheffield.
- 10 Department of Health. *The health of the nation: a strategy for health in England*. London: HMSO, 1993.
- 11 Cottler LB, Zipp JF, Robins LN, Spitznagel EL. Difficult to recruit respondents and their effect on prevalence estimates in and epidemiological survey. *Am J Epidemiol* 1987;125:329–39.
- 12 OPCS. *Standard occupational classification*. London: HMSO, 1991.
- 13 Leyden R, Gobel F, Killoran A. *Health and life-style surveys: towards a common approach*. London: HEA/OPCS, 1990.
- 14 Sengupta C. *The reliability and validity of information from a postal lifestyle survey*. Nottingham: University of Nottingham, 1993, BMedSci Thesis.